A practical query-by-humming system for a large music database

Full text Pdf (1.05 MB)

Source **International Multimedia Conference** archive
Proceedings of the eighth ACM international conference on Multimedia table of contents
Marina del Rey, California, United States
Pages: 333 - 342
Year of Publication: 2000
ISBN:1-58113-198-4

Authors
Naoko Kosugi NTT Laboratories, 1-1, Hikarinooka, Yokosuka-shi, Kanagawa, 239-0847, Japan
Yuichi Nishihara NTT Laboratories, 1-1, Hikarinooka, Yokosuka-shi, Kanagawa, 239-0847, Japan
Tetsuo Sakata NTT Laboratories, 1-1, Hikarinooka, Yokosuka-shi, Kanagawa, 239-0847, Japan
Masashi Yamamuro NTT Laboratories, 1-1, Hikarinooka, Yokosuka-shi, Kanagawa, 239-0847, Japan
Kazuhiko Kushima NTT Laboratories, 1-1, Hikarinooka, Yokosuka-shi, Kanagawa, 239-0847, Japan

Sponsors
SIGCHI: ACM Special Interest Group on Computer-Human Interaction
SIGCOMM: ACM Special Interest Group on Data Communication
SIGIR: ACM Special Interest Group on Information Retrieval
SIGMULTIMEDIA: ACM Special Interest Group on Multimedia
SIGOPS: ACM Special Interest Group on Operating Systems
SIGGRAPH: ACM Special Interest Group on Computer Graphics and Interactive Techniques
SIGMIS: ACM Special Interest Group on Management Information Systems

Publisher ACM Press New York, NY, USA

↑ ABSTRACT

A music retrieval system that accepts hummed tunes as queries is described in this paper. This system uses similarity retrieval because a hummed tune may contain errors. The retrieval result is a list of song names ranked according to the closeness of the match. Our ultimate goal is that the correct song should be first on the list. This means that eventually our system's similarity retrieval should allow for only one correct answer.

The most significant improvement our system has over general query-by-humming systems is that all processing of musical information is done based on beats instead of notes. This type of query processing is robust against queries generated from erroneous input. In addition, acoustic information is transcribed and converted into relative intervals and is used for making feature vectors. This increases the resolution of the retrieval system compared with other general systems, which use only pitch direction information.

The database currently holds over 10,000 songs, and the retrieval time is at most one second. This level of performance is mainly achieved through the use of indices for retrieval. In this paper, we also report on the results of music analyses of the songs in the database. Based on these results, new technologies for improving retrieval accuracy, such as partial feature vectors and or'ed retrieval among multiple search keys, are proposed. The effectiveness of these technologies is evaluated quantitatively, and it is found that the retrieval accuracy increases by more than 20% compared with the previous system [9]. Practical user interfaces for the system are also described.

↑ REFERENCES

Note: OCR errors may be found in this Reference List extracted from the full text article. ACM has opted to expose the complete List rather than only correct and linked references.

1 Steven Blackburn , David DeRoure, A tool for content based navigation of music, Proceedings of the sixth ACM international conference on Multimedia, p.361-368, September 13-16, 1998, Bristol, United Kingdom

2 K. Curtis, N. Taniguchi, J. Nakagawa, and M. Yamamuro. A comprehensive image similarity retrieval system that

utilizes multiple feature vectors in high dimensional space. In Proceedings of International Conference on Inforraation, Communication and Signal Processing, pages 180-184, September 1997.

3 J. J. Dubnowski, R. W. Schafer, and L. R. Rabiner. Real-Time Digital Hardware Pitch Detector. IEEE Trans. on Acoustics, Speech, and Signal Processing, ASSP-24(1):2-8, February 1976.

4 Jonathan Foote, An overview of audio information retrieval, Multimedia Systems, v.7 n.1, p.2-10, Jan. 1999

5 Jonathan Foote, Visualizing music and audio using self-similarity, Proceedings of the seventh ACM international conference on Multimedia (Part 1), p.77-80, October 30-November 05, 1999, Orlando, Florida, United States

6 J. T. Foote . Content-Based Retrieval of Music and Audio. In Proc. SPIE, vol3229, pages 138-147, 1997.

7 Asif Ghias , Jonathan Logan , David Chamberlin , Brian C. Smith, Query by humming: musical information retrieval in an audio database, Proceedings of the third ACM international conference on Multimedia, p.231-236, November 05-09, 1995, San Francisco, California, United States

8 Naoko Kosugi , Yuïchi Nishihara , Seiichi Kon'ya , Masashi Yamamuro , Kazuhiko Kushima, Let's search for songs by humming!, Proceedings of the seventh ACM international conference on Multimedia (Part 2), p.194, October 30-November 05, 1999, Orlando, Florida, United States

9 N. Kosugi, Y. Nishihara, S. Kon'ya, M. Yamamuro, and K. Kushima. Music Retrieval by Humming. In Proceedings of PACRIM'99, pages 404-407. IEEE, August 1999.

10 W. Y. Ma, B. S. Manjunath, Y. Luo, Y. Deng, and X. Sun. NETRA: A Content-Based Image Retrieval System. http://maya.ece.uesb.edu/Netra/.

11 R. J. MeNab, L. A. Smith, D. Bainbridge, and I. H. Witten. The New Zealand Digital Library MELody inDEX. http://www.dlib.org/dlib/may97/meldex/OSwritten.html, May 1997.

12 Muscle Fish LLC. http://www.musclefish.eom/.

13 Y. Nishihara, N. Kosugi, S. Kon'ya, and M. Yamamuro. Humming Query System Using Normalized Time Scale. In Proceedings of CODAS'99, March 1999.

14 Pierre-Yves Rolland , Gailius Raškinis , Jean-Gabriel Ganascia, Musical content-based retrieval: an overview of the Melodiscov approach and system, Proceedings of the seventh ACM international conference on Multimedia (Part 1), p.81-84, October 30-November 05, 1999, Orlando, Florida, United States

15 John R. Smith , Chung-Sheng Li, Image classification and querying using composite region templates, Computer Vision and Image Understanding, v.75 n.1-2, p.165-174, July/Aug. 1999

16 WILDCAT CANYON SOFTWARE. AUTOSCORE. http://www.wildcat.com/Pages/AutoscoreMain.htm.

17 Noburou Taniguchi , Masashi Yamamuro, Multiple Inverted Array Structure for Similar Image Retrieval, Proceedings of the IEEE International Conference on Multimedia Computing and Systems, p.160, June 28-July 01, 1998

18 Alexandra Uitdenbgerd , Justin Zobel, Melodic matching techniques for large music databases, Proceedings of the seventh ACM international conference on Multimedia (Part 1), p.57-66, October 30-November 05, 1999, Orlando, Florida, United States

19 Sun Wu , Udi Manber, Fast text searching: allowing errors, Communications of the ACM, v.35 n.10, p.83-91, Oct. 1992

20 M. Yamamuro, K. Knshima, H. Kimoto, H. Akama, S. Kon'ya, J. Nakagawa, K. Mii, N. Taniguchi, and K. Curtis. ExSight - Multimedia Information Retrieval System. In 2Oth Annual Pacific Telecommunications Conference, PTC'98 Proceedings, pages 734-739, 1998.

21 Atsuo Yoshitaka , Tadao Ichikawa, A Survey on Content-Based Retrieval for Multimedia Databases, IEEE Transactions on Knowledge and Data Engineering, v.11 n.1, p.81-93, January 1999

↑ **CITINGS** ***14***

Liu Wenyin , Zheng Chen , Minjing Li , Hongjiang Zhang, A media agent for automatically building a personalized semantic index, Journal of the American Society for Information Science and Technology, v.52 n.10, p.853-855, August 2001

Naoko Kosugi , Go Nishimura , Junji Teramoto , Kazuyoshi Mii , Makoto Onizuka , Seiichi Kon'ya , Akira Kojima , Ryoji Kataoka , Takashi Honishi , Kazuhiko Kushima, Content-based retrieval applications on a common database management system, Proceedings of the ninth ACM international conference on Multimedia, September 30-October 05, 2001, Ottawa, Canada

Yazhong Feng , Yueting Zhuang , Yunhe Pan, A hierarchical approach: query large music database by acoustic input, Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval, August 11-15, 2002, Tampere, Finland

Andreas Rauber , Elias Pampalk , Dieter Merkl, Content-based music indexing and organization, Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval, August 11-15, 2002, Tampere, Finland

Naoko Kosugi , Yasushi Sakurai , Masashi Morimoto, SoundCompass: a practical query-by-humming system; normalization of scalable and shiftable time-series data and effective subsequence generation, Proceedings of the 2004 ACM SIGMOD international conference on Management of data, June 13-18, 2004, Paris, France

Cheng Yang, Peer-to-peer architecture for content-based music retrieval on acoustic data, Proceedings of the 12th international conference on World Wide Web, May 20-24, 2003, Budapest, Hungary

Cheng Yang, Efficient acoustic index for music retrieval with various degrees of similarity, Proceedings of the tenth ACM international conference on Multimedia, December 01-06, 2002, Juan-les-Pins, France

Richard L. Kline , Ephraim P. Glinert, Approximate matching algorithms for music information retrieval using vocal input, Proceedings of the eleventh ACM international conference on Multimedia, November 02-08, 2003, Berkeley, CA, USA

Liu Wenyin , Zheng Chen , Fan Lin , Hongjiang Zhang , Wei-Ying Ma, Ubiquitous media agents: a framework for managing personally accumulated multimedia files, Multimedia Systems, v.9 n.2, p.144-156, August 2003

Chih-Chin Liu , Chuan-Sung Huang, A singer identification technique for content-based classification of MP3 music objects, Proceedings of the eleventh international conference on Information and knowledge management, November 04-09, 2002, McLean, Virginia, USA

Elias Pampalk , Andreas Rauber , Dieter Merkl, Content-based organization and visualization of music archives, Proceedings of the tenth ACM international conference on Multimedia, December 01-06, 2002, Juan-les-Pins, France

Jyh-Shing Roger Jang , Hong-Ru Lee, Hierarchical filtering method for content-based music retrieval via acoustic input, Proceedings of the ninth ACM international conference on Multimedia, September 30-October 05, 2001, Ottawa, Canada

Shyamala Doraisamy , Stefan Rüger, Robust Polyphonic Music Retrieval with *N*-grams, Journal of Intelligent Information Systems, v.21 n.1, p.53-70, July 2003

Yunyue Zhu , Dennis Shasha, Warping indexes with envelope transforms for query by humming, Proceedings of the 2003 ACM SIGMOD international conference on Management of data, June 09-12, 2003, San Diego, California

↑ INDEX TERMS

Primary Classification:

H. Information Systems

↳ **H.5** INFORMATION INTERFACES AND PRESENTATION (I.7)

Additional Classification:

H. Information Systems

↳ **H.2** DATABASE MANAGEMENT

↳ **H.2.4** Systems

↳ **Subjects:** Multimedia databases

- **H.3** INFORMATION STORAGE AND RETRIEVAL
  - **H.3.3** Information Search and Retrieval
    - **Subjects:** Retrieval models
- **H.5** INFORMATION INTERFACES AND PRESENTATION (I.7)
  - **H.5.1** Multimedia Information Systems
    - **Subjects:** Audio input/output
  - **H.5.2** User Interfaces (D.2.2, H.1.2, I.3.6)
    - **Subjects:** Graphical user interfaces (GUI)

General Terms:
Design, Management, Measurement, Performance, Theory

Collaborative Colleagues:

Naoko Kosugi:	Takashi Honishi	Makoto Onizuka
	Ryoji Kataoka	Tetsuo Sakata
	Akira Kojima	Yasushi Sakurai
	Seiichi Kon'ya	Yoshiharu Suga
	Kazuhiko Kushima	Junji Teramoto
	Kazuyoshi Mii	Mario Tokoro
	Atsushi Mitsuzawa	Masashi Yamamuro
	Masashi Morimoto	
	Yuichi Nishihara	
	Go Nishimura	
Kazuhiko Kushima:	Takehito Abe	Takahiko Murayama
	Hiroki Akama	Yuichi Nishihara
	Hiroshi Fujii	Go Nishimura
	Takashi Honishi	Makoto Onizuka
	Shoichiro Ishigaki	Tetsuo Sakata
	Ryoji Kataoka	Junji Teramoto
	Akira Kojima	Masashi Yamamuro
	Seiichi Kon'ya	
	Naoko Kosugi	
	Kazuyoshi Mii	
Yuichi Nishihara:	Seiichi Kon'ya	
	Naoko Kosugi	
	Kazuhiko Kushima	
	Tetsuo Sakata	
	Masashi Yamamuro	
Tetsuo Sakata:	Naoko Kosugi	
	Kazuhiko Kushima	
	Yuichi Nishihara	
	Masashi Yamamuro	
Masashi Yamamuro:	Hiroki Akama	Yuichi Nishihara
	Takeharu Eda	Makoto Onizuka
	Atsushi Katayama	Tetsuo Sakata
	Mitsuru Kawashimo	Noboru Sonehara
	Seiichi Kon'ya	Gengo Suzuki
	Naoko Kosugi	Noburou Taniguchi
	Kazuhiko Kushima	
	Masaru Nakagawa	
	Takao Nakamura	
	Hidekazu Nakawatase	

Peer to Peer - Readers of this Article have also read:

- Constructing reality **Proceedings of the 11th annual international conference on Systems documentation**
  Douglas A. Powell , Norman R. Ball , Mansel W. Griffiths
- M^4: a metamodel for data preprocessing **Proceedings of the 4th ACM international workshop on Data warehousing and OLAP**
  Anca Vaduva , Jörg-Uwe Kietz , Regina Zücker
- Data structures for quadtree approximation and compression **Communications of the ACM** 28, 9

Hanan Samet

- A hierarchical single-key-lock access control using the Chinese remainder theorem **Proceedings of the 1992 ACM/SIGAPP Symposium on Applied computing**
  Kim S. Lee , Huizhu Lu , D. D. Fisher

- An intelligent component database for behavioral synthesis **Proceedings of the 27th ACM/IEEE conference on Design automation**
  Gwo-Dong Chen , Daniel D. Gajski